Volatilidad

Gabriel V. Montes-Rojas

¿Por qué es importante la volatilidad?

La volatilidad es el estudio de la varianza conditional (o desviación estándar) de un proceso

$$r_t = \mu_t + a_t \; (media)$$

$$\sigma_t^2 = Var(a_t|F_{t-1}) = E(r_t - \mu_t)^2 \; (varianza)$$

Aquí $\{\mu_t\}$ *puede ser un proceso ARMA o un modelo de regresión* $\mu_t = \beta_0 + \beta_1 x_t$, $\{a_t\}$ *son los shocks.*

La idea es modelar σ_t^2 *como un proceso AR o MA. La volatilidad puede no ser constante (heteroscedasticidad general). En muchos casos se observa clusters de volatilidad: períodos de alta (baja) volatilidad son seguidos de peródos de alta (baja) volatilidad .*
La volatilidad no es directamente observable al igual que los shocks. Entonces corresponde a una variable latente. (Los modelos de alta frequencia, donde hay muchas observaciones en el día intraday, pueden estimar la volatilidad para períodos agregados.)

¿Por qué es importante la volatilidad?

- La volatilidad es un factor importante para options trading (Black-Scholes).
- Value at Risk (VaR) o Expected Shortfall (ES).
- Estimar la volatilidad mejora la eficiencia de los estimadores.
- Leverage effects: los cambios en los precios están negativamente correlacionados con cambios en la volatilidad (riesgo vs. retorno esperado).
- Efectos de la nueva información: alta volatilidad es observada antes de que se hagan anuncios.

Supuestos de Gauss-Markov

Consideremos una regresión simple:

$$y_i = \beta_0 + \beta_1 x_i + u_i, \ \ i = 1, 2, ..., n$$

- Supuesto 5: Homoscedasticidad $Var(u_i|\mathbf{X}) = \sigma^2, \ \ i = 1, 2, ..., n$
- Teorema de Gauss-Markov Theorem: Bajo los Supuestos 1-5, el estimador MCO es el mejor estimador lineal insesgado (best linear unbiased estimators, BLUE). Nota: Mejor significa de menor varianza.

La heteroscedasticidad se define como $Var(u_i|\mathbf{X}_i) = \sigma_i^2$ para $i = 1, 2, ..., n$. Significa que la varianza del error no es constante.

Inferencia vs. Insesgadez

- El principal problema con la heteroscedasticidad es que $Var(\hat{\beta}_1)$ no es la que sería estimada bajo los supuestos de Gauss-Markov. ¡Esto falla! ¿Por qué? Tomemos como ejemplo un modelo de regresión simple con homoscedasticidad:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Ahora una expresión más general es

$$Var(\hat{\beta}_1) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sigma_i^2}{SST_x^2}$$

  donde $\sigma_i^2 = Var(u_i | x_i)$

- Entonces, **la inferencia no es válida debido a que se usaría errores estándar incorrectos.**
- Sin embargo, los estimadores MCO siguen siendo **insesgados** y **consistentes**. ¿Por qué?

Autoregressive Conditional Heteroskedasticity (ARCH)

Engle, R. F. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflations," Econometrica, 50, 987–1007.

ARCH(m)

$$a_t = \sigma_t \, \epsilon_t \; (\textit{mean equation})$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \ldots + \alpha_m a_{t-m}^2 \; (\textit{volatility equation})$$

- donde $\{\epsilon_t\}$ es una secuencia de variables aleatorias iid con media 0 y varianza 1.
- $\alpha_0 > 0$ y $\alpha_i \geq 0$ para $i = 1, 2, \ldots, m$. Los coeficientes tienen que satisfacer ciertas condiciones de regularidad para que la varianza no condicional de a_t sea finita.
- En la práctica a_t se asume que sigue una distribución normal o t-Student.

ARCH(1)

ARCH(1)

$$a_t = \sigma_t \, \epsilon_t \ (mean\ equation)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 \ (volatility\ equation)$$

- $E(a_t) = E[E(a_t|F_{t-1})] = E[\sigma_t E(\epsilon_t)] = 0.$
- $Var(a_t) = E(a_t^2) = E[E(a_t^2|F_{t-1})] = \alpha_0 + \alpha_1 E(a_{t-1}^2).$
- Si el proceso es estacionario, entonces $Var(a_t) = \alpha_0/(1-\alpha_1)$. Entonces necesitamos $0 \leq \alpha_1 < 1$.
- Asumiendo normalidad [kurtosis=3] $E(a_t^4|F_{t-1}) = 3E(a_t^2|F_{t-1}) = 3(\alpha_0 + \alpha_1 a_{t-1}^2)^2$. Entonces $E(a_t^4) = E[E(a_t^4|F_{t-1})] = 3E[\alpha_0^2 + 2\alpha_0\alpha_1 a_{t-1}^2 + \alpha_1^2 a_{t-1}^4]$. Usando estacionariedad y $m_4 = E(a_t^4)$, $m_4 = 3(\alpha_0^2 + 2\alpha_0\alpha_1 Var(a_t) + \alpha_1^2 m_4)$,

$$m_4 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_2^2)}, 0 \leq \alpha_1^2 < \frac{1}{3}$$

- El modelo impone restricciones en los valores de los parámetros para la existencia de los parámetros de simetría (skewness) y curtosis (kurtosis).
- El modelo asume que los shocks positivos y negativos tienen el mismo efecto sobre la volatilidad porque depende de los cuadrados de los shocks previos. En la práctica sin embargo se sabe que las series financieras responden diferente a shocks positivos que a shocks negativos.
- El modelo ARCH no provee ninguna información aerca de la(s) causa(s) en las variaciones de la volatilidad. Sólo provee un mecanismo para describir la varianza condicional.
- El modelo ARCH tiende a sobre-predecir la volatilidad porque responde a shocks grandes y aislados.

Construcción del modelo

Pasos a seguir para construir un modelo ARCH:

1. Especificar la ecuación de la media usando modelos ARMA.
2. Usar los residuos de la media para testear por efectos ARCH.
3. Seleccionar el orden del modelo ARCH (m).
4. Especificar la volatilidad del modelo y chequear que sean estadísticamente significativos.
5. Chequear el modelo especificado y refinar si es necesario.

Contraste para efectos ARCH

2 alternativas:
(1) Usar los estadísticos de Ljung-Box $Q(m)$ y las funciones de autocorrelación (ACF, autocorrelation function) a los residuos de la ecuación de la media $\{a_t^2\}$.

- Especificar la ecuación de la media y estimar los residuos $\hat{a}_t$, computar $\hat{a}_t^2$
- Aplicar ACF a $\{\hat{a}_t^2\}$
- Para testear por ARCH(m) usar $Q(m)$ sobre $\{\hat{a}_t^2\}$: La hipótesis nula es que los primeros m rezagos lags de la ACF de a_t^2 son 0.

(2) Lagrange multiplier (LM) test de Engle (1982)

- Este contraste es equivalente al contraste F para $\alpha_i = 0$, $(i = 1, ..., m)$ en la regresión lineal

$$a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + ... + +\alpha_m a_{t-m}^2 + e_t, \; t = m+1, ..., T$$

- $H_0 : \alpha_1 = ... = \alpha_m = 0$

Seleccionar el orden de ARCH

- Definir $\eta_t = a_t^2 - \sigma_t^2$. Se puede demostrar que $\{\eta_t\}$ no esta autocorrelacionada y tiene media 0.
- Reescribir $a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \ldots + +\alpha_m a_{t-m}^2 + \eta_t$
- Esto es un modelo $AR(m)$ para los shocks al cuadrado
- Usar la función de autocorrelación parcial (PACF, partial autocorrelation function) para determinar el orden del modelo AR.

Estimar ARCH

- Varias funciones de verosimilitud se usan para la estimación de ARCH, dependiendo de la distribución de ϵ_t: normal, t-Student, GED
- Si consideramos la función de verosimilitud normal:

$$f(a_1, ..., a_T|\alpha) = f(a_T|F_{T-1})f(a_{T-1}|F_{T-2})f(a_{m+1}|F_m) \times f(a_m, ..., a_1|\alpha)$$

$$= \prod_{t=m+1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} exp(-\frac{a_t}{2\sigma_t^2}) \times f(a_m, ..., a_1|\alpha)$$

- Las condiciones iniciales son difíciles de manejar $f(a_m, ..., a_1|\alpha)$. En general se consideran supuestos o se ignora.
- Log-likelihood

$$\ell(a_{m+1}, ..., a_T|\alpha, a_m, ..., a_1) = -\sum_{t=m+1}^{T} \left[\frac{1}{2} ln(\sigma_t^2) + \frac{1}{2}\frac{a_t^2}{\sigma_t^2}\right]$$

ARCH: STATA

Para estimar ARCH:

```
tsset TIME
arch YVAR, arch(m)
```

Diferentes distribuciones de las innovaciones se pueden considerar en la opción `distribution()`. Opciones: normal, t, ged

ARCH se puede combinar con modelos ARMA en la ecuación de la media. Por ejemplo, un ARCH(m) con ARMA(p,q) en la media:

```
arch YVAR, arch(m) ar(p) ma(q)
```

ARCH: STATA

Para contrastar por efectos ARCH. LM test de Engle (1982)

```
tsset TIME
reg YVAR
estat archlm, lags(m)
```

Alternativamente luego de la estimación ,

```
test [ARCH]L1.arch [ARCH]L2.arch ...  [ARCH]Lm.arch
```

Para seleccionar el orden

```
reg YVAR
predict res, resid
gen res2=res*res
pac res2
```

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

Bollerslev, T. (1986), "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, 31, 307–327.

- Consideremos la diferencia entre un modelo AR y uno MA.
- Una diferencia similar distingue entre ARCH y GARCH.
- GARCH produce un modelo más parsimonioso que ARCH.
- GARCH también reproduce los cluster de volatilidad mejor que ARCH (al igual que AR vs. MA para la media).

Generalized Autoregressive Conditional Heteroskedasticity (GARCH)

GARCH(m,s)

$$a_t = \sigma_t \, \epsilon_t \; (\textit{mean equation})$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{m} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2 \; (\textit{volatility equation})$$

- donde $\{\epsilon_t\}$ es una secuencia de shocks iid con media 0 y varianza 1.
- $\alpha_0 > 0$, $\alpha_i \geq 0$ para $i = 1, 2, ..., m$, $\beta_j \geq 0$ para $j = 1, 2, ..., s$. Los coeficientes tienen que satisfacer ciertas condiciones de regularidad para asegurar que la varianza no condicional de a_t sea finita.
- α_i y β_j se definen respectivamente como los parámetros ARCH y GARCH.

GARCH(1,1)

GARCH(1,1)

$$a_t = \sigma_t \, \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- Definir $\eta_t = a_t^2 - \sigma_t^2$ tal que $\sigma_t^2 = a_t^2 - \eta_t$
- Notar que $\{\eta_t\}$ es una secuencia de diferencias martingala (martingale difference sequence) con $E(\eta_t) = 0$ y $cov(\eta_t, \eta_{t-j}) = 0, j > 0$.
- Entonces, $a_t^2 = \alpha_0 + (\alpha_1 + \beta_1) a_{t-1}^2 + \eta_t - \beta_1 \eta_{t-1}$, que es como un proceso ARMA(1,1).
- Así, $E(a_t^2) = \frac{\alpha_0}{1-\alpha_1-\beta_1}$, y necesitamos $1 - \alpha_1 - \beta_1 > 0$.
- Asumiendo estacionariedad se llega a

$$\frac{E(a_t^4)}{[E(a_t^2)]^2} = \frac{3(1-(\alpha_1+\beta_1)^2)}{1-(\alpha_1+\beta_1)^2 - 2\alpha_1^2} > 3$$

Entonces lo modelos GARCH generan una distribución no condicional con colas más pesadas.

GARCH: STATA

Para estimar modelos GARCH:
```
tsset TIME
arch y, arch(m) garch(s)
```

Para contrastar por GARCH luego de estimarlo,
```
test [ARCH]L1.garch [ARCH]L2.garch ...  [ARCH]Ls.garch
```

Para contrastar por ARCH y GARCH luego de estimarlo,
```
test [ARCH]L1.arch [ARCH]L2.arch ...  [ARCH]Lm.arch [ARCH]L1.garch
[ARCH]L2.garch ...  [ARCH]Ls.garch
```

¿Cómo construir la volatilidad estimada con modelos ARCH y GARCH?

- Recordemos que la voltilidad es no observable, es una variable latente.
- Los modelos ARCH/GARCH nos permiten construir la volatilidad.
- Luego de estimar un modelo:

  ```
  predict sigma2hat, variance
  ```

  Esto produce una serie estimada de la volatilidad dentro de la muestra $\{\hat{\sigma}_t^2\}$, llamada `sigma2hat`

  ```
  line sigma2hat TIME
  ```

¿Cómo proyectar con modelos ARCH/GARCH?

- Consideremos un modelo simple GARCH(1,1) con AR(1) en la media: `arch YVAR, arch(1) garch(1) ar(1)`
- ¿Cómo construir $\hat{y}_{t+1}$ y $\hat{\sigma}^2_{t+1}$?
- Hay que usar las fórmulas recursivamente.
- Para la proyección un período hacia adelante (1-step ahead forecast), $\sigma^2_h(1) = E(\sigma^2_{h+1}|F_h) = \alpha_0 + \alpha_1 a^2_h + \beta_1 \sigma^2_h$ (donde a_h and σ^2_h se conocen en el tiempo h)
- Para proyectar muchos períodos $a^2_t = \sigma^2_t \epsilon^2_t$, y de la ecuación de volatilidad $\sigma^2_{t+1} = \alpha_0 + (\alpha_1 + \beta_1)\sigma^2_t + \alpha_1 \sigma^2_t(\epsilon^2_t - 1)$. Entonces,

$$\begin{aligned}\sigma^2_h(2) &= \alpha_0 + (\alpha_1 + \beta_1)\sigma^2_h(1) + \alpha_1 \sigma^2_h(1) E(\epsilon^2_{h+1} - 1|F_h) \\ &= \alpha_0 + (\alpha_1 + \beta_1)\sigma^2_h(1)\end{aligned}$$

  using $E(\epsilon^2_{h+1} - 1|F_h) = 0$
- Entonces proyectar ℓ períodos se computa como

$$\sigma^2_h(\ell) = \frac{\alpha_0[1 - (\alpha_1 + \beta_1)^{\ell-1}]}{1 - \alpha_1 - \beta_1} + (\alpha_1 + \beta_1)^{\ell-1}\sigma^2_h(1)$$

- $\sigma^2_h(\ell) \to \frac{\alpha_0}{1-\alpha_1-\beta_1}$, *as* $\ell \to \infty$

¿Cómo proyectar con modelos ARCH/GARCH?

`qui summ`
`gl N1=r(N)+1`
`set obs $N1` [para incrementar la muestra +1]
`summ TIME`
`replace TIME=r(max)+1 in $N1` [para incrementar la variable de tiempo, TIME, +1]
`tsset TIME` [hay que especificar nuevamente]

`predict yhat` [predicción dentro de la muestra $\hat{y}_t$]
`predict uhat, resid` [predicción de los residuos dentro de la muestra $\hat{u}_t$]
`gen uhat2=uhat*uhat`
`predict sigma2hat, variance`

```
replace yhatfcast=_b[_cons]+[ARMA]_b[L1.ar]*L1.YVAR
replace sigma2hatfcast=[ARCH]_b[_cons]+[ARCH]_b[L1.arch]*L1.uhat2 +
[ARCH]_b[L1.garch]*L1.sigma2hat
```

Otros ejemplos en:
`http://www.learneconometrics.com/class/5263/notes/arch.pdf`

GARCH in the mean (GARCHM)

Engel, R.F., Lilien, D.M. y Robins, R.P. (1987), "Estimating time varying risk premia in the term structure: The Arch-M model," Econometrica, 55(2), 391–407.

$$r_t = \mu + c\sigma_t^2 + a_t$$

$$a_t = \sigma_t \, \epsilon_t \; (\textit{mean equation})$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \; (\textit{volatility equation})$$

- En finanzas el retorno de un activo depende del riesgo/volatilidad.
- en este modelo el parámetro c se llama *risk premium* (si $c > 0$ significa que la volatilidad incremwenta el retorno de los activos).
- La autocorrelación serial en r_t se explica por el risk premium (efficient market hypothesis).
- GARCHM debe usarse cuando la teoría soporta un trade-off entre riesgo y retorno.
- En STATA:
  `arch y, archm arch(m) garch(s) archmlags(h)`
  donde h es el número de lags de σ^2 que aparecen en la ecuación de la media.

Exponential GARCH (EGARCH)

Nelson, D.B. (1991), "Conditional heteroskedasticity in assets returns: A new approach," Econometrica, 59(2), 347-370.

- La idea es permitir efectos asimétricos de los shocks en la volatilidad.
- Consideremos la función asimétrica $g(\epsilon_t) = \theta\epsilon_t + \gamma[|\epsilon_t| - E(|\epsilon_t|)]$
  Entonces,

$$g(\epsilon_t) = \begin{array}{l} (\theta+\gamma)\epsilon_t - \gamma E(|\epsilon_t|) \text{ } if \text{ } \epsilon_t \geq 0 \\ (\theta-\gamma)\epsilon_t - \gamma E(|\epsilon_t|) \text{ } if \text{ } \epsilon_t < 0 \end{array}$$

- Este modelo es similar a un modelo de umbrales con umbral=0.
  $E(|\epsilon_t|)$ varía de acuerdo a la distribución que se asuma de ϵ. Si $\epsilon \sim N(0,1)$
  Entonces $E(|\epsilon_t|) = \sqrt{2/\pi}$

Exponential GARCH (EGARCH)

- El modelo usa el logaritmo de la varianza condicional para imponer menos restricciones de no negatividad.
- EGARCH(m,s)

$$a_t = \sigma_t \, \epsilon_t \; (mean \; equation)$$

$$ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + ... + \beta_m B^s}{1 - \alpha_1 B - ... - \alpha_m B^m} g(\epsilon_t) \; (volatility \; equation)$$

donde B es el operador de lags.

- An EGARCH(1,1) model is

$$a_t = \sigma_t \, \epsilon_t \; (mean \; equation)$$

$$(1 - \alpha B) ln(\sigma_t^2) = (1 - \alpha)\alpha_0 + g(\epsilon_t) \; (volatility \; equation)$$

- Para estimar en STATA: `arch y, earch(m) egarch(s)`

Integrated GARCH (IGARCH)

Si el polinomio AR de la ecuación de volatilidad tiene una raíz unitaria, entonces usamos el modelo IGARCH. En este caso el efecto de los shocks del pasado $\eta_{t-i} = a_{t-i}^2 - \sigma_{t-i}^2$, $i > 0$ en a_t^2 son persistentes.

IGARCH(1,1)

$$a_t = \sigma_t \, \epsilon_t \; (\textit{mean equation})$$

$$\sigma_t^2 = \alpha_0 + \beta_1 \sigma_{t-1}^2 + (1-\beta_1) a_{t-1}^2 \; (\textit{volatility equation})$$

- donde $\{\epsilon_t\}$ es una secuencia de ruido blanco.
- $1 > \beta_1 > 0$

Multivariate volatility

- Generalize the univariate volatility models the multivariate case
- By multivariate volatility, we mean the conditional covariance matrix of multiple asset returns
- Multivariate volatilities have many important financial applications. They play an important role in portfolio selection and asset allocation, and they can be used to compute the value at risk of a financial position consisting of multiple assets

Multivariate volatility

Consider a multivariate return series $\{\mathbf{r}_t\}$

$$\mathbf{r}_t = \mu_t + \mathbf{a}_t$$

where $\mu_t = E(\mathbf{r}_t|\mathbf{F}_{t-1})$ is the conditional expectation and $\mathbf{r}_t$ is the shock or innovation. A typical application is an ARMA structure. Now define

$$\Sigma_t = Cov(\mathbf{a}_t|\mathbf{F}_{t-1})$$

- Multivariate volatility models refer to the evolution of Σ_t
- Curse of dimentionality: k(k + 1)/2 quantities in Σ_t for a k-dimensional return series
- We need to guarantee it is positive definite (it is a variance!)

Exponentially weighted estimate

Given the innovations $\mathbf{F}_{t-1} = \{\mathbf{a}_1, ..., \mathbf{a}_{t-1}\}$ the (unconditional) covariance innovation can be estimated by

$$\hat{\Sigma}_t = \frac{1}{t-1} \sum_{j=1}^{t-1} \mathbf{a}_j \mathbf{a}_j'$$

- Similar idea to White (1980) heteroskedasticity-robust estimator of the variance-covariance in OLS
- Note that it uses past information on correlations based on $\mathbf{a}_j \mathbf{a}_j'$

Exponentially weighted estimate

To allow for a time-varying matrix

$$\hat{\Sigma}_t = \frac{1-\lambda}{1-\lambda^{t-1}} \sum_{j=1}^{t-1} \lambda^{j-1} \mathbf{a}_j \mathbf{a}_j'$$

where $0 < \lambda < 1$ and the weights $\frac{(1-\lambda)\lambda^{j-1}}{1-\lambda^{t-1}}$ sum to one.
For a sufficiently large t we have $\lambda^{t-1} \approx 0$, then

$$\hat{\Sigma}_t \approx (1-\lambda)\mathbf{a}_j \mathbf{a}_j' + \lambda \hat{\Sigma}_{t-1}$$

(exponentially weighted moving-average (EWMA) estimate of the covariance matrix)

Multivariate GARCH

- Multivariate GARCH models allow the conditional covariance matrix of the dependent variables to follow a flexible dynamic structure. Bollerslev, Engle, and Wooldridge (1988); Bollerslev, Engle, and Nelson (1994); Bauwens, Laurent, and Rombouts (2006); and Silvennoinen and Terasvirta (2009) discuss the general multivariate GARCH model and many special cases. There is a plethora of multivariate GARCH models because the unrestricted multivariate GARCH model is too general to fit to data and there are trade-offs in creating restricted models.
- Same idea to VAR (vector autoregressive model) but also applied to volatility

Diagonal Vectorization (VEC) Model

Bollerslev, Engle, and Wooldridge (1988) generalize the exponentially weighted moving-average approach to propose a multivariate GARCH.

DVEC(m, s) model

$$\Sigma_t = A_0 + \sum_{i=1}^{m} A_i \odot (\mathbf{a}_j \mathbf{a}_j') + \sum_{j=1}^{s} B_j \odot \Sigma_{j-1}$$

where m and s are positive integers, A_i and B_j are symmetric matrices, and $\odot$ is the Hadamard product element-by-element multiplication

Diagonal Vectorization (VEC) Model

DVEC(1, 1) model

$$\begin{bmatrix} \sigma_{11,t} & \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} A_{11,0} & \\ A_{21,0} & A_{22,0} \end{bmatrix} +$$

$$\begin{bmatrix} A_{11,1} & \\ A_{21,1} & A_{22,1} \end{bmatrix} \odot \begin{bmatrix} a^2_{1,t-1} & \\ a_{1,t-1}a_{2,t-1} & a^2_{2,t-1} \end{bmatrix} +$$

$$\begin{bmatrix} B_{11,1} & \\ B_{21,1} & B_{22,1} \end{bmatrix} \odot \begin{bmatrix} \sigma_{1,t-1} & \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix}$$

Note: there are 9 parameters to estimate
$A_{11,0}, A_{21,0}, A_{22,0}, A_{11,1}, A_{21,1}, A_{22,1}, B_{11,1}, B_{21,1}, B_{22,1}$

Diagonal Vectorization (VEC) Model

DVEC(1, 1) model

$$\sigma^2_{11,t} = A_{11,0} + A_{11,1}a^2_{1,t-1} + B_{11,1}\sigma_{11,t-1}$$
$$\sigma^2_{21,t} = A_{21,0} + A_{21,1}a_{1,t-1}a_{2,t-1} + B_{11,1}\sigma_{21,t-1}$$
$$\sigma^2_{22,t} = A_{22,0} + A_{22,1}a^2_{2,t-1} + B_{22,1}\sigma_{22,t-1}$$

That is, each element of a DVEC model follows a GARCH(1,1)-type model. However, it may not produce a positive-definite covariance matrix. Furthermore, the model does not allow for dynamic dependence between volatility series.

Diagonal Vectorization (VEC) Model

How to implement DVEC in STATA? (Need STATA 11 or more....)

Suppose we have 2 assets YVAR1 and YVAR2, and we have in mind an ARMA model or some model with covariates with variables XVAR1 and XVAR2. Note that XVAR1 and XVAR2 could be the same.

```
dvech (YVAR1=XVAR1) (YVAR2=XVAR2), arch(m) garch(s)
```

This could be further simplified to

```
dvech (YVAR1 YVAR2=XVAR1 XVAR2), arch(m) garch(s)
```

if both YVAR1 and YVAR2 depend on the same XVARs

BEKK Model

To guarantee the positive-definite constraint, Engle and Kroner (1995) propose the Baba-Engle-Kraft-Kroner (BEKK) model,

$$\Sigma_t = AA' + \sum_{i=1}^{m} A_i(\mathbf{a}_i\mathbf{a}_i')A_i' + \sum_{j=1}^{s} B_j\Sigma_{j-1}B_j'$$

where A is a lower triangular matrix and A_i and B_j are k $\times$ k matrices. Based on the symmetric parameterization of the model, Σ_t is almost surely positive definite provided that AA' is positive definite.

Reparametrization

- Can we do something to simplify the estimation of Σ?
- Use of correlation. Separate correlation and variance
- Cholesky decomposition

Use of correlations

$$\Sigma_t \equiv [\sigma_{ij,t}] = D_t \rho_t D_t,$$

where ρ_t is the conditional correlation matrix of $\mathbf{a}_t$, and D_t is a $k \times k$ diagonal matrix consisting of the conditional standard deviations of $\mathbf{a}_t$, that is $D_t = diag\{\sqrt{\sigma_{11,t}}, ..., \sigma_{kk,t}\}$

With STATA you can also implement Multivariate GARCH models

`http://www.stata.com/stata12/multivariate-garch/`